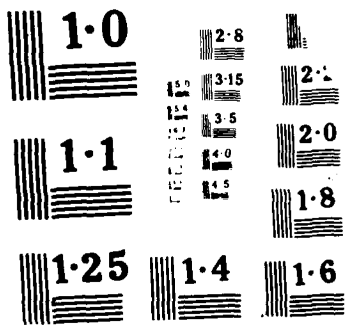
AD-A191 509 DELAY-DOPPLER RADAR IMAGING(U) WASHINGTON UNIV ST LOUIS 1/1
MO ELECTRONIC SYSTEMS AND SIGNALS RESEARCH LAB
D L SNYDER 30 NOV 87 ESSRL-88-1 N00014-86-K-0370

UNCLASSIFIED F/G 17/9 NL


1·0
2·8
3·15
3·5
2·0
1·1
4·0
4·5
1·8
1·25
1·4
1·6

Washington

WASHINGTON·UNIVERSITY·IN·ST·LOUIS

AD-A191 509

(4)

DTIC FILE COPY

DELAY-DOPPLER RADAR IMAGING

Semi-Annual Progress Report No. 3
O.N.R. Contract N00014-86-K-0370
Period: 1 June 1987 - 30 November 1987

DTIC
ELECTE
FEB 1 9 1988
S D
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Principal Investigator: Donald L. Snyder

ESSRL-88-1

ELECTRONIC SYSTEMS
AND SIGNALS
RESEARCH LABORATORY

Department of Electrical Engineering
Campus Box 1127
One Brookings Drive
Washington University

88 2 2 005

DELAY-DOPPLER IMAGING-RADAR

Semi-Annual Progress Report No. 3

Office of Naval Research Contract Number N00014-86-K-0370

Period Covered: 1 June 1987 - 30 November 1987

Principal Investigator:

Donald L. Snyder
Director, Electronic Systems and Signals Research Laboratory
Washington University
Campus Box 1127
One Brookings Drive
St. Louis, Missouri 63130

Scientific Program Director:

Dr. Rabinder Madan
Office of Naval Research
Code 1114SE
800 North Quincy Street
Arlington, Virginia 22217-5000

Accesion For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By per ltr.
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

DISTRIBUTION

	copies
Mr. John W. Michalski Office of Naval Research Resident Representative Federal Building, Room 286 536 South Clark Street Chicago, Illinois 60605-1588	1
Dr. Rabinder N. Madan Office of Naval Research 800 N. Quincy Street Code 1114SE Arlington, Virginia 22217-5000	1
Director Naval Research Laboratory Attn.: Code 2627 Washington, DC 20375	1
Defense Technical Information Center Building 5 Cameron Station Alexandria, Virginia 22314	12
Mr. Harper J. Whitehouse Naval Ocean Systems Center Code 7402 San Diego, California 92152	1
Dr. James Fienup Environmental Research Institute of Michigan P.O. Box 8618 Ann Arbor, MI 48107	1
Dr. Kenneth Senne M.I.T. Lincoln Laboratory Lexington, MA 02173	1

Table of Contents

1 Introduction

This third semi-annual progress report contains a summary of work accomplished on O. N. R. contract number N00014-86-K-0370, *Delay-Doppler Radar-Imaging*, during the period from 1 June 1987 to 30 November 1987. This six-month period is the last of three making up the 18 month period of this contract. A one-year continuation contract entitled *High Resolution Radar-Imaging* has been awarded for the period 1 December 1987 to 30 November 1988.

The goal of this project is to formulate and investigate new approaches for forming images of radar targets from spotlight-mode, delay-doppler measurements. These measurements could be acquired with a high-resolution radar-imaging system operating with an optical- or radio-frequency carrier. Two approaches are under study. The first is motivated by an image-reconstruction algorithm used in radionuclide imaging called the *confidence-weighted algorithm*; here, we will refer to this approach as the *chirp-rate modulation approach*. The second approach is based on more fundamental principles which starts with a mathematical model that accurately describes the physics of an imaging radar-system and then uses statistical-estimation theory with this model to derive processing algorithms; we will refer to this as the *estimation-theory approach*.

Spotlight-mode high-resolution radar-imaging relies upon the relative motion between the transmitter, target, and receiver. The target is illuminated by a series of transmitted pulses. The return for each pulse is a superposition of reflections from various locations on the target, with each location affecting the pulse by introducing a time delay, doppler shift, and reflectance gain. The returns are processed to produce an image of the target. Two types of images are possible. One is a map in delay/doppler or range/cross-range coordinates of the target's complex-valued reflectivity, which indicates the amplitude-gain applied to the incident radar-pulse by each location on the target. The other is a map in the same coordinates of the target's scattering function, which indicates the power-gain applied at each location on the target. We are studying approaches for producing both types of images.

A common approach discussed in the literature and implemented in practice is to use the same transmitted pulse for each illumination of the target. The returns are processed using a

two-dimensional Fourier transform to produce the target's image in delay/doppler or range/cross-range coordinates [1,2]. One of our goals is to compare images produced in this standard way with those produced using the alternative approaches we are developing.

Bernfeld [3] appears to be the first to introduce the idea for radar imaging of modifying the pulse shape on successive illuminations of the target. We are using Bernfeld's idea in the chirp-rate modulation approach. With this, the FM chirp-rate of each pulse is varied so that the angles made by the ambiguity functions in the delay/doppler plane are caused to vary over the full range of angles between 0° and 180°. Use is then made of the fact that, on the average, the square-law envelope-detected output of a receiver consisting of a bank of filters matched to doppler shifted versions of a transmitted pulse is the two-dimensional convolution of the ambiguity function of the pulse with the scattering function of the target [4], an output we call the delay/doppler power function. Given the delay/doppler power functions for each illumination, the target's scattering function can be determined using the confidence-weighted algorithm [4].

Statistical models for radar echoes from spatially extended targets that are rough compared to a wavelength of the radar carrier are given by Van Trees [5] for microwave frequencies and Shapiro, Capron, and Harney [6] for optical frequencies. In the estimation-theory approach, we are using these models with statistical estimation-theory to develop new algorithms for producing high resolution images of radar targets from delay/doppler data [7].

Work accomplished during the reporting period is summarized in the following section.

2 Summary of Work Accomplished

2.1 Estimation-Theory Approach to Imaging

During this reporting period, reference [7, see Appendix 1 for a preprint] was revised and submitted to the IEEE Transactions on Information Theory for review and publication. This manuscript contains a discussion of the mathematical model we are using to describe radar returns from diffuse radar targets. Equations are derived for the maximum-likelihood estimate of the target's reflectivity process and its scattering function.

In this project, the complex envelope of the radar-echo data is described by:

$$r(t) = \sqrt{2E_T}\int_{-\infty}^{\infty} s(t-\tau)b\left(t-\frac{1}{2}\tau,\tau\right)d\tau + w(t),$$

where E_T is the transmitted energy, $s(t)$ is the complex envelope of the transmitted pulse-sequence, $w(t)$ is broad band Gaussian noise, and $b(t,\tau)$ is the target's reflectivity at time t for all the reflecting patches on the target at two-way delay τ. This model describes the radar data as a superposition of echoes from all reflecting patches on the target plus an additive noise. The reflectivity $b(t,\tau)$ is assumed to be a zero-mean, wide-sense stationary, uncorrelated-scatterers (WSSUS) Gaussian process with a scattering function $S(f,\tau)$ in doppler f and delay τ coordinates; this model for the reflectivity is motivated in references [5] and [6]. The scattering function is the power-density spectrum of the reflectivity process $b(t,\tau)$ at two-way delay τ. The problem we address in [7] is that of forming the maximum-likelihood estimate of the scattering function based on data $r(t)$. This problem is addressed in a general framework, for example with the form of the transmitted pulse left arbitrary, so that the fundamental equations describing the estimate can be obtained for a wide variety of potentially interesting special cases. The estimate satisfies a nonlinear integral-equation; we describe an iterative approach for solving the equation numerically. The unique feature of our approach is that we have addressed the problem of radar imaging by starting with a basic statistical model for the return data and then deriving an image-formation algorithm using estimation theory. We expect that improved target images will be obtained in those situations where the reflectivity model is an accurate description of reality.

At the present time, we are beginning to investigate scattering-function estimates in computer simulations, and we are attempting to extend the model of the target's reflectivity to include specular or glint components in the echo signal in addition to the diffuse component contained in the model of [7]. We also wish to incorporate constraints, implied by targets of interest, into the image-formation process.

2.2 Chirp-Rate Modulation Approach to Imaging

The focus of our research on the chirp-rate modulation approach during the reporting period has been the analytical investigation of a linear architecture based on a bank of bandpass

matched-filters. The motivation for this resulted from a concern that the nonlinear architecture being used in our investigations, a bank of bandpass matched-filters followed by a square-law envelope detector, results in a loss of coherence in the processing of data from successive target illuminations and that this loss could result in degraded target images.

We consider the above model for return data $r(t)$, but temporarily without the additive noise $w(t)$. By writing the reflectivity process $b(t,\tau)$ in terms of its Fourier transform $c(f_d,\tau)$,

$$b(t,\tau) = \int_{-\infty}^{\infty} c(f_d,\tau) e^{j2\pi f_d t} df_d,$$

we can express the return data according to

$$r(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(f_d,\tau) s(t-\tau) e^{j2\pi f_d\left(t-\frac{1}{2}\tau\right)} df_d d\tau.$$

The bank of bandpass matched filters produces the complex-valued delay/doppler image $q(t,f)$ defined by

$$\begin{aligned} q(t,f) &= \int_{-\infty}^{\infty} r(u) s^*(u-t) e^{-j2\pi f(u-t)} du \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(f_d,\tau)\phi(\tau-t, f_d-f) e^{-j\pi(f\tau-f_d t)} df_d d\tau e^{-j\pi f t}, \end{aligned}$$

where

$$\phi(\tau, f_d) = \int_{-\infty}^{\infty} s\left(z-\frac{1}{2}\tau\right) s^*\left(z+\frac{1}{2}\tau\right) e^{j2\pi f_d z} dz$$

is the complex-valued time-frequency autocorrelation function of the transmitted pulse (also sometimes called the complex-valued ambiguity function). Thus, the data at the output of the bank of matched filters are a linear functional of $c(f_d,\tau)$ and the time-frequency autocorrelation function. The goal is to solve the inverse problem of determining $c(f_d,\tau)$ from a collection of filter outputs $q_i(\tau,f)$, where the ith output results from the ith transmitted pulse in the chirp-rate modulated sequence. One potentially interesting observation that can be made is that if the time-frequency autocorrelation function is slowly varying (ideally, a constant) over the

extent of the target image $c(f_d,\tau)$ in the delay/doppler plane, then $q(t,\tau)$ is related to the two-dimensional Fourier transform of $c(f_d,\tau)$, and the reflectivity image may then be formed by inverse Fourier-transformation in the absence of noise.

In the progress report submitted for the first six month period of this contract, there was a table listing various parameter values being used in computer simulation experiments with the confidence-weighted algorithm. An update of this table is contained in Appendix 2.

3 References

1. D. L. Mensa, *High Resolution Radar Imaging*, Artech House, Dedham, MA, 1981.

2. D. R. Wehner, *High Resolution Radar*, Artech House, Dedham, MA, 1987.

3. M. Bernfeld, "CHIRP Doppler Radar," Proc. IEEE, Vol. 72, pp. 540-541, April 1984.

4. D. L. Snyder, H. J. Whitehouse, J. T. Wohlschlaeger, and R. C. Lewis, "A New Approach to Radar/Sonar Imaging," Proc. 1986 SPIE Conference on Advanced Algorithms and Architectures, Vol. 696, pp. 134-139, 1986.

5. H. L. Van Trees, *Estimation, Detection, and Modulation Theory, Part III*, John Wiley and Sons, NY, 1968.

6. J. Shapiro, B. A. Capron, and R. C. Harney, "Imaging and Target Detection with a Heterodyne-Reception Optical Radar," Applied Optics, Vol. 20, No. 19, pp. 3292-3313, October 1981.

7. D. L. Snyder, J. A. O'Sullivan, and M. I. Miller, "The Use of Maximum-Likelihood Estimation for Forming Images of Diffuse Radar-Targets from Delay-Doppler Data," IEEE Transactions on Information Theory, in review. Appendix 1 contains a preprint.

4 Appendices

4.1 Appendix 1. Preprint of Reference [7].

THE USE OF MAXIMUM-LIKELIHOOD ESTIMATION FOR FORMING IMAGES OF DIFFUSE RADAR-TARGETS FROM DELAY-DOPPLER DATA†

Donald L. Snyder
Joseph A. O'Sullivan
and
Michael I. Miller

Electronic Systems and Signals Research Laboratory
Department of Electrical Engineering
Washington University
St. Louis, MO 63130

ABSTRACT

In this paper, we present a new approach to high-resolution radar imaging. The starting point is a model of the radar echo-signal based on the physics governing radar reflections. This model has been used several times in the past for describing radar targets that are rough compared to the wavelength of the transmitted radiation. Without specifying precisely what the transmitted signal is, we derive a general estimation-based procedure for obtaining images. After discretizing the model, the radar imaging problem reduces to the task of estimating discretized second-order statistics of the reflectance process of the target. A maximum-likelihood estimate of these statistics is obtained as the limit point of an expectation-maximization algorithm.

† This work was supported by the Office of Naval Research under contract N00014-86-K0370.

1. Introduction

Radar systems can be used to produce high-resolution images of a reflecting target. This is accomplished by illuminating the target with a series of pulses and observing the return echos. Each patch on the target introduces a certain amount of propagation delay and doppler shift to a pulse it reflects, the amount depending on the range and velocity of the patch relative to the radar transmitter and receiver. The beam-width of the radar antenna relative to the size of the target is an important consideration. Images can be produced by scanning a narrowly focused beam over the target in some type of raster pattern and then displaying the received power in delay and doppler or, equivalently, range and cross-range coordinates. Images can also be formed by illuminating the entire target in *spotlight mode* with a wide, relatively unfocused beam. The received signal for each illumination is then a complicated superposition of the echos received from all the patches that make up the extended surface of the radar target. Our concern will be with forming images of rotating, rough targets using a spotlight-mode radar.

We will denote the complex envelope of the signal transmitted by the radar as $(2E_T)^{1/2}s_T(t)$, where E_T is the transmitted energy, and $s_T(t)$ is normalized to unit energy. The particular form of this signal will not need to be specified. The expressions we obtain for producing an image can then be specialized for any signal of choice, including the stepped-frequency and wideband chirp waveforms used in practice, as discussed by Wehner [1] and Mensa [2] and the chirp-rate modulated waveforms discussed by Bernfeld [3] and Snyder, Whitehouse, Wohlschlaeger and Lewis [4]. When specializing $s_T(t)$, it should be kept in mind that this represents the entire sequence of transmitted pulses that illuminate the target.

Following Walker [5], consider a small, nonfluctuating reflector that is rotating about an axis at the rate of f_r revolutions each second, as shown in Fig. 1. The distance from the radar transmitter/receiver to the axis of rotation is R_0, and the distance to the reflector at time t is given approximately by

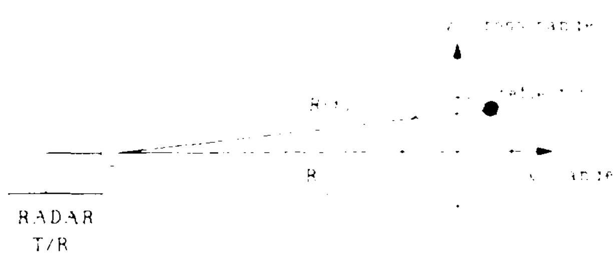


Figure 1. Radar geometry for a small rotating reflector

$$R(t) \doteq R_0 + x_0 \cos(2\pi f_r t) + y_0 \sin(2\pi f_r t).$$

provided $R_0 >> (x_0^2 + y_0^2)^{1/2}$. Then, the radar echo-signal $s_R(t)$ received following an illumination by $s_T(t)$ will be of the form

$$s_R(t) = \sqrt{2E_T}\, s_T(t-\tau) b$$

where $\tau = 2R(t)/c$ is the two-way propagation delay to the reflector, with c being the propagation velocity. The quantity b is a complex-valued scale factor which models the strength of the received echo. This scale factor will be called the *reflectivity*. It includes the effects of inverse square-law attenuation experienced by the propagating radiation and, importantly, the properties of the reflector that are significant in the electromagnetic scattering interaction, including its shape, size and surface properties. More generally, the reflectivity can vary with time because the aspect and, therefore, the scattering interaction with the reflector will vary as it rotates. If, as discussed by Walker [5], the radar data $s_R(t)$ are examined over a small interval of time, then the delay τ and doppler shift f_D can be approximated by

$$\tau \doteq 2c^{-1}(R_0 + x_0),$$

and

$$f_D = \frac{2}{\lambda}\frac{dR}{dt} \doteq \frac{2}{\lambda} y_0 2\pi f_r,$$

where λ is the wavelength at the carrier frequency of the radar. Thus, the reflectivity b, range x_0, cross-range y_0, relative to the coordinate axis, can be determined from the amplitude, delay, and doppler information contained in the radar data. Extracting this information permits an image of the reflector to be formed by displaying $|b|$ or $|b|^2$ at the appropriate location in range and cross-range coordinates. The maximum delay and doppler shift are determined by the distance $(x_0^2 + y_0^2)^{1/2}$ of the reflector from the coordinate center about which it rotates and the rotation rate f_r; more generally, the extent of a reflector in delay and doppler is determined by the physical extent of the reflector and the rotation rate.

Now consider a spatially extended target that is rotating. A patch on the surface with a two-way delay in the interval $[\tau, \tau+\Delta\tau)$ reflects a signal that is incident on the patch at time t

with a reflectance strength $b(t,\tau)\Delta\tau$. Consequently, the complex envelope of the received echo-signal $s_R(t)$ following the illumination of the target by $s_T(t)$ is given by the following superposition of returns from reflecting patches at all the two-way delays τ :

$$s_R(t) = \sqrt{2E_T}\int_{-\infty}^{+\infty} s_T(t-\tau)\, b\left(t-\frac{1}{2}\tau,\tau\right)d\tau . \tag{1}$$

The total received signal, $r(t)$, is also assumed to be corrupted by an independent, additive noise,

$$r(t) = s_R(t) + w(t), \tag{2}$$

where $w(t)$ is a complex-valued, white Gaussian-process with a mean of zero and a covariance function defined by

$$\mathrm{E}[w(t)w^*(t')] = N_0\delta(t-t'), \tag{3}$$

where " * " denotes complex conjugation. We refer to $b(t,\tau)$ as the *reflectance process.* This is a complex-valued random process.

There are two images which may be displayed as the result of processing $r(t)$ with an estimation procedure. One is an estimate of the reflectance process $b(t,\tau)$ itself and the other is an estimate of the covariance or, equivalently, the spectral density of this process. These may be regarded as conditional first-order and a second-order statistics of the reflectivity process, respectively, in terms of the radar data (2).

If $b(f,\tau)$ is deterministic, define $c(t,\tau)$ to be its Fourier transform in the t variable,

$$c(f,\tau) = \int_{-\infty}^{+\infty} b(t,\tau)e^{-j2\pi ft}dt . \tag{4}$$

The function $c(f,\tau)$ then contains information about the target in delay, τ, and doppler, f, coordinates. An image of the target is obtained by placing the magnitude or squared magnitude of this function into delay and doppler bins. We refer to this as the *reflectance image.* This transform can be obtained in a variety of ways, depending on the signal $s_T(t)$ selected to illuminate the target. For the stepped-frequency signals used in practice, the usual approach consists of two operations described by Wehner [1]. The first is to place the data into delay (or, range) bins by separately Fourier transforming N sample values of the received signal acquired

for each transmitted group of N stepped-frequency pulses. The resulting delay-binned data are placed in the rows of an $N \times N$ matrix, where each row contains the transformed data from one pulse-group. In the second operation, the columns of this matrix are Fourier transformed to obtain a doppler (or, cross-range) profile at each delay. The resulting two-dimensional array is intended to be a discrete version of $c(f,\tau)$ in delay (range) and doppler (cross-range) coordinates. This processing based on two-dimensional Fourier transforms is derived using a strictly deterministic analysis and so does not account for statistical properties of the reflectance process or for noise that may be present. A similar processing is employed for the linear FM-chirp signals also used in practice for radar imaging [1, 2].

For situations in which the target's surface is rough compared to the wavelength at the carrier frequency, $b(t,\tau)$ may be taken to be a complex-valued Gaussian random process, as discussed by Shapiro, Capron, and Harney [6] for radar systems operating at laser frequencies and Van Trees [7] at microwave frequencies. If there are no glint or specular components in the echo, then this is a zero mean process with covariance

$$E[b(t,\tau)b^*(t',\tau')] = K(t-t',\tau)\delta(\tau-\tau'). \tag{5}$$

The delta function in this expression results from postulating that each reflecting patch introduces an uncorrelated contribution to the echo signal. That the function $K(t-t',\tau)$ depends only on the difference of t and t', and not on t and t' separately, results from postulating that the reflectance process is wide-sense stationary for each delay. A reflectance process with these properties is said by Van Trees [7] to possess wide-sense stationary, uncorrelated scatterers (WSSUS). Assuming that the reflectance process has these properties, the delay-doppler data associated with the radar target may be obtained from the Fourier transform of $K(t,\tau)$ in the t variable,

$$S(f,\tau) = \int_{-\infty}^{\infty} K(t,\tau)\exp(-j2\pi ft)dt. \tag{6}$$

The function $S(f,\tau)$ is called the *scattering function* of the target and, as a function of f, is the power-density spectrum of the reflectance process at each delay τ. $S(f,\tau)\Delta f \Delta\tau$ is the mean-square strength (or power) of the reflectance of all patches on the target having a doppler

shift in the interval $[f,f+\Delta f)$ and a delay in the interval $[\tau,\tau+\Delta\tau)$. The scattering function may be viewed in delay and doppler coordinates as an image of the target. We call this the *scattering function image*.

Our approach to forming radar images will be to use maximum-likelihood methods with the data model in (2) to estimate the scattering function. We will also obtain an estimate of the reflectance process. Model based approaches that use statistical estimation-theory techniques to derive image-formation algorithms appear less frequently in the large literature about radar imaging than do the deterministic approaches outlined above. One example is that of Frost, Stiles, Shanmugan, and Holtzman [8], who use a multiplicative model and Wiener-filtering techniques. The approach we will describe below differs in that the model (2) we adopt of the echo signal is more complicated than a simple multiplicative one and depends explicitly on the transmitted waveform through a spatial integration over the reflecting target. We also do not restrict the processing to be linear; in particular, we show that the processing of the received data for producing the maximum-likelihood estimate of the scattering function and a corresponding estimate of the reflectance process is not linear. An approach for estimating scattering functions of spread channels is given by Gaarder [9], who cites earlier work on the subject by Green [10], Kailath [11], Gallager [12], Hagfors [13,14], Price [15], Levin [16], Abraham [17], Sifford [18], and Reiffen [19]. Gaarder assumes a specific processing architecture in the form of a cascade of a linear filter, square-law envelope detector, and another linear filter, and he claims that this processing is either more general than or equivalent to those of most previous authors. Our approach differs in that no particular processing is assumed in advance; rather, we derive the processing to produce the estimates, starting from a model for the received data and applying recent results in maximum-likelihood estimation. The processing which results is quite distinct from that discussed by Gaardner.

For our new approach to radar imaging, we adopt the WSSUS model of a diffuse radar target described by Shapiro, Capron, and Harney [6] and Van Trees [7]. We treat both the reflectance process and its second-order statistic, the scattering function, as unknown quantities. The iterative approach we develop for forming images yields the maximum-likelihood estimate of the scattering function and, simultaneously, the conditional-mean estimate of the reflectance

process based on statistics which are consistent with the estimated scattering function. Thus, both of the quantities treated separately in other imaging schemes are produced simultaneously with our new approach, which is a unique and important aspect of our approach.

We will develop a necessary condition, called the *trace condition*, which the maximum-likelihood estimate of the target's scattering function must satisfy. This equation appears to be very hard to solve analytically. As a consequence, we reformulate the imaging problem using the concept of incomplete-complete data spaces and then use the expectation-maximization algorithm of Dempster, Laird, and Rubin [20] to derive an iterative algorithm for producing the maximum-likelihood estimate of the scattering function. The technique we use to accomplish this parallels that described by Miller and Snyder [21] for power-spectrum estimation and extends their work to include indirect measurements of the process whose spectrum is sought; the process is now measured following the linear transformation and additive noise seen in (2). As shown by Turmon and Miller [22], this is a high-resolution approach to spectrum estimation which results in estimated spectra with less bias and mean-square error than other recently developed approaches discussed in the literature. We expect that similar improvements will be seen in radar images of scintillating, diffuse targets when this new technique is used.

2. Maximum-Likelihood Imaging for the Incomplete-Data Model

For reasons that will become evident in the next section, we term the data $r(t)$ in (2) the *incomplete data* for the radar-imaging problem. The model given in the Introduction for these data is as the sum of the radar echo-signal $s_R(t)$ of (1) and an independent, white noise-process $w(t)$. We may, therefore, state the problem of imaging a diffuse radar-target as that of estimating the scattering function $S(f,\tau)$ or, equivalently, the covariance function $K(t,\tau)$ given radar-return data $\{r(t),\ T_i \leq t \leq T_f\}$ on an observation interval (T_i,T_f). In this section, we first discretize the model for the incomplete data, and then we derive and discuss a necessary condition, called the *trace condition*, which the maximum-likelihood estimate of the discretized version of $K(t,\tau)$ must satisfy.

discrete model

In anticipation of using discrete-time processing of radar data to produce images, we now state the discrete version of our model as follows. We are given N samples of the complex-valued radar-data corresponding to (2),

$$r(n) = s_R(n) + w(n), \qquad n = 0, 1, \ldots, N-1, \tag{7}$$

where $w(n)$ is a white Gaussian-sequence with zero mean and covariance

$$\mathrm{E}[w(n)w^*(n')] = N_0 \delta_{n,n'}, \tag{8}$$

and where the signal samples corresponding to (1) are given by

$$s_R(n) = \sqrt{2E_T} \sum_{i=-\infty}^{+\infty} s_T(n,i) b(n,i), \qquad n = 0, 1, \ldots, N-1. \tag{9}$$

In this expression, we define $s_T(n,i)$ and $b(n,i)$ in terms of the transmitted signal and the reflectance process, respectively, according to:

$$s_T(n,i) = s_T(n\Delta t - i\Delta\tau), \tag{10}$$

and

$$b(n,i) = b\left(n\Delta t - \frac{1}{2} i\Delta\tau, i\Delta\tau\right)\Delta t. \tag{11}$$

where Δt and $\Delta \tau$ are the sampling intervals adopted in the discretization in time and delay, respectively. We assume that the target has a *finite extent in range*, so $b(n,i)$ and, therefore also, terms forming the sum in (9) are zero for i outside the I_R (here, the subscript R denotes *range*) values $m, m+1, \ldots, m+I_R-1$ starting from the minimum two-way delay corresponding to m. This discrete reflectance is a Gaussian sequence with zero mean and covariance given by

$$E[b(n,i)b^*(n',i')] = K(n-n',i)\delta_{i,i'}, \tag{12}$$

where $\delta_{i,i'}$ is the Kronecker delta-function. The discrete scattering function $S(f,i)$ is the Fourier transform of $K(n,i)$ in the n variable,

$$S(\nu,i) = \sum_{n=-\infty}^{+\infty} K(n,i)\exp(-j2\pi\nu n). \tag{13}$$

The imaging problem for the discrete model is to estimate $S(\nu,i)$, or equivalently the covariance function $K(n,i)$, for all frequencies ν spanning the target in doppler, and for all delays i spanning the target in delay, given the radar data $\{r(n),\ n=0,\ 1,\ \ldots,\ N-1\}$.

matrix model

These discrete equations may conveniently be written in matrix form as follows. Define r to be the received-signal vector of dimension N,

$$r = \begin{pmatrix} r(0) \\ r(1) \\ \cdot \\ \cdot \\ \cdot \\ r(N-1) \end{pmatrix} = s_R + w, \tag{14}$$

where the N-dimensional vectors, s_R and w, are given by

$$s_R = \begin{pmatrix} s_R(0) \\ s_R(1) \\ \cdot \\ \cdot \\ s_R(N-1) \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w(0) \\ w(1) \\ \cdot \\ \cdot \\ w(N-1) \end{pmatrix}. \tag{15}$$

Also, define S^* as the $NI_R \times N$ rectangular matrix expressed in column-block form in terms of I_R separate $N \times N$ matrices according to

$$S^* = \begin{pmatrix} S_0 \\ S_1 \\ \cdot \\ \cdot \\ \cdot \\ S_{I_R - 1} \end{pmatrix}, \tag{16}$$

where S_j is an $N \times N$ diagonal matrix containing sample values of the complex envelope of the transmitted signal $s_T(t)$,

$$S_j = \begin{pmatrix} s_T(0, m+j) & 0 & 0 & \cdot & \cdot & 0 \\ 0 & s_T(1, m+j) & 0 & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & s_T(N-1, m+j) \end{pmatrix}. \tag{17}$$

Further, define the reflectance vector b of dimension NI_R in the column-block form of I_R vectors according to

$$b = \begin{pmatrix} b(0) \\ b(1) \\ \cdot \\ \cdot \\ \cdot \\ b\ I_R - 1 \end{pmatrix}, \tag{18}$$

where each $b(i)$ is a vector of dimension N,

$$b(i) = \begin{pmatrix} b(0, m+i) \\ b(1, m+i) \\ \cdot \\ \cdot \\ \cdot \\ b(N-1, m+i) \end{pmatrix}. \tag{19}$$

Using (9) and these definitions, we can now express the N-dimensional signal vector s_R of (14) and (15) as

$$s_R = \sqrt{2E_T}\, S^+ b, \tag{20}$$

where superscript "+" denotes the Hermitian-transpose operation.

In terms of these defined matrices, the received vector has zero mean and covariance

$$\begin{aligned} K_r = \mathrm{E}\{rr^+\} &= \mathrm{E}\{s_R s_R^+\} + \mathrm{E}\{uu^+\} \\ &= 2E_T S^+ \mathrm{E}\{bb^+\} S + N_0 I. \end{aligned} \tag{21}$$

Then, since

$$\mathrm{E}\{b(i)b^+(j)\} = K(i)\delta_{i,j}, \tag{22}$$

where $K(i)$ is the Hermitian-symmetric Toeplitz-matrix

$$K(i) = \begin{pmatrix} K(0,m+i) & K^*(1,m+i) & \cdot & \cdot & K^*(N-1,m+i) \\ K(1,m+i) & K(0,m+i) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ K(N-1,m+i) & \cdot & \cdot & \cdot & K(0,m+i) \end{pmatrix}, \tag{23}$$

it follows from (21) that the covariance K_r of r is given by

$$K_r = 2E_T S^+ K S + N_0 I, \tag{24}$$

where K is the block-diagonal $NI_R \times NI_R$-dimensional matrix defined by

$$K = \begin{pmatrix} K(0) & 0 & 0 & \cdot & \cdot & 0 \\ 0 & K(1) & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & K(I_R-1) \end{pmatrix}. \tag{25}$$

The i-*th* diagonal block $K(i)$ of K is the covariance matrix of the reflectance process at the i-*th* delay bin.

the estimation problem

We will use the following definition.

Definition: Let **K** denote the set of all $NI_R \times NI_R$ block-diagonal matrices (25) with each block $K(i)$ an $N \times N$ Hermitian-symmetric Toeplitz matrix (23). Let $\Omega \subset \mathbf{K}$ be a specified convex subset of **K**. Any matrix $K \in \Omega$ is termed *admissible*. A variational matrix $\delta K \in \mathbf{K}$ is called an *admissible variation* of K for a fixed $K \in \Omega$ if there exists an $\alpha > 0$ such that $K + \beta\delta K \in \Omega$ for all β satisfying $0 \le \beta \le \alpha$.

The problem is to form an admissible estimate of the covariance matrix K of (25) given the data vector r of (14). The radar image then viewed is the discrete scattering function, an estimate of which may be obtained from the estimate of K by use of (13).

In the definition above, the constraint that K be in Ω is used to obtain a "reasonable" setup of the problem. Here, we assume that the scattering function $S(f,\tau)$ in (6) is only nonzero for frequencies f satisfying $|f| \le f_{max}$ for some finite upper frequency f_{max} and for all delays i. This is equivalent to the assumption of a target of finite cross section and rotation rate. The discrete-time scattering function $S(\nu,i)$ of (13) is then a periodic function of ν consisting of a sum of shifted replicas of $S(f,\tau)$ scaled in amplitude by $1/\Delta t$ and in frequency by Δt, where Δt is the time between samples of $r(t)$. The replicas of $S(f,\tau)$ are centered at every integer on the ν scale. In order to guarantee that there is no aliasing, assume that the sampling rate, $1/\Delta t$, of $r(t)$ satisfies the Nyquist condition, $1/\Delta t > 2f_{max}$. Then, $S(\nu,i)$ will be nonzero between $-1/2$ and $+1/2$ only for ν satisfying $|\nu| \le \nu_{max} = f_{max}\Delta t$. The output of our algorithm is $S(\nu,i)$ discretized in frequency ν. For a resolution having at least I_{CR} (here, the subscript CR denotes *cross range*) samples in the frequency range $-\nu_{max} \le \nu \le \nu_{max}$, a total of

$$P \ge \frac{I_{CR}}{2\Delta t f_{max}}$$

samples of ν between $-1/2$ and $+1/2$ are required.

The model for the incomplete data r of (14) is that r is normally distributed with zero mean and covariance specified in (24). Given the incomplete data, we wish to estimate the covariance K of the reflectance process, as defined in (25). To do this, we adopt the maximum-likelihood method of statistics, which selects K to maximize the incomplete-data loglikelihood

$$L_{id}(K) = -\frac{1}{2}\ln\left(\det\left(2E_T S^* K S + N_0 I\right)\right) - \frac{1}{2} r^* \left(2E_T S^* K S + N_0 I\right)^{-1} r \,. \tag{26}$$

where the maximization is subject to the constraint that K be an admissible matrix.

Lemma 1. A necessary condition for an admissible K in the interior of Ω to be a local maximum of $L_{\mathrm{id}}(K)$ over all $K \in \Omega$ is

$$\mathrm{Tr}\Big((2E_T S^* K S + N_0 I)^{-1}(rr^* - 2E_T S^* K S - N_0 I)(2E_T S^* K S + N_0 I)^{-1} S^* \delta K S\Big) = 0, \tag{27}$$

for all admissible variations δK.

The proof of Lemma 1 in the Appendix is based on the fact that the necessary condition for an admissible K to maximize $L_{\mathrm{id}}(K)$ is that for all admissible variations δK, there holds

$$\lim_{\alpha \downarrow 0^+} \frac{L_{\mathrm{id}}(K + \alpha \delta K) - L_{\mathrm{id}}(K)}{\alpha} \le 0. \tag{28}$$

We call (27) the *trace condition*. Burg, Luenberger, and Wenger [23] have studied an equivalent problem of Toeplitz-constrained covariance-estimation and have derived the trace condition using a different approach.

If $\Omega = \mathbf{K}$, there are NI_{R} unknowns in K. Since $\delta K \in \mathbf{K}$, there are NI_{R} parameters in δK that can be varied for this case. These variations in the trace condition generate NI_{R} equations in the unknown elements of K. Thus, in principle, the trace condition produces enough equations to determine the unconstrained maximum-likelihood estimate K. However, the equations are complicated due to the inverse matrices appearing in (27), so it does not appear to be feasible to determine K directly from the trace condition. This motivates our development of the iterative approach in the next section. A sequence of estimates that increase the likelihood at each iteration stage is identified, and we demonstrate that limit points of the iteration satisfy the trace condition.

The trace condition is only a necessary condition which the estimate K must satisfy. For it to be sufficient as well, the second derivative must be negative along all admissible variational directions δK.

Lemma 2. Sufficient conditions for an admissible matrix K to be a local maximum of $L_{\mathrm{id}}(K)$ are that, first, the trace condition (27) is satisfied for all admissible variations δK and, second, that there holds

$$Tr\left(K_r^{-1} S^* \delta K S K_r^{-1}\left(2E_T S^* K S + N_0 I - 2rr^*\right) K_r^{-1} S^* \delta K S\right) < 0 \tag{29}$$

for all admissible variations δK

The proof of Lemma 2 is in the Appendix. Equation (29) is just the second derivative of $L_{\mathrm{id}}(K)$ in the direction δK.

3. Maximum-Likelihood Imaging for the Incomplete/Complete Data Model

That the trace condition (27) cannot be solved directly for the maximum-likelihood estimate of K motivates the indirect approach we now take of embedding the imaging problem in a larger, seemingly more difficult problem. The result will be an iterative algorithm which when implemented produces a sequence of admissible matrices $K^{(0)}, K^{(1)}, \ldots, K^{(k)}, \ldots$ having the property that the corresponding sequence of incomplete-data loglikelihoods $L_{id}[K^{(0)}]$, $L_{id}[K^{(1)}]$, ..., $L_{id}[K^{(k)}]$, ... is nondecreasing at each stage.

Fuhrmann and Miller [24] have recently shown that maximum-likelihood estimates of Toeplitz-constrained covariances which are positive definite do not always exist when given only one data vector r. A necessary and sufficient condition for the likelihood function to be unbounded, and therefore for no maximum-likelihood estimate to exist, is that there be a singular Toeplitz matrix with the data in its range space. For our imaging problem, this condition is that there exists an admissible K with

$$2E_T S^* K S + N_0 I$$

singular such that

$$r = \left(2E_T S^* K S + N_0 I\right)\alpha$$

for some complex-valued vector α. In fact, without constraining K further than being Toeplitz, a sufficient condition that a singular estimate for K be obtained is that $N_0 = 0$ and there exists a singular K with r in the range space of $2E_T S^+ KS$. The argument for this mirrors that of Fuhrmann and Miller in [24, Theorem 1] but is applied to the complete data loglikelihood given below in equation (A7) in the Appendix. Furhmann and Miller also showed that even if the true covariance had eigenvalues bounded from above and below, the probability that there exists a singular Toeplitz matrix with the data in its range can be very close to one. By restricting the search to Toeplitz matrices with circulant extensions, they were able to show that the probability a singular circulant Toeplitz matrix has the data in its range space is zero. Thus, in order for maximum-likelihood estimates to be nonsingular with probability one for all nonnegative values of N_0, we may restrict the class of admissible Toeplitz matrices to be those with circulant extensions of period P, where P is equal to or greater than the number N of data samples available, $P \geq N$. What we envision in adopting this constraint is that for each delay i, the N sample values of the reflectance $b(n,i)$, $n = 0, 1, \ldots, N-1$, are from a stationary process

that is periodic with period P, where P could be some large but finite value; a lower bound on P in terms of a desired cross-range resolution is discussed above. These N sample values of the reflectance enter the incomplete data r according to (14) and (20). By using the expectation-maximization algorithm of Dempster, Laird, and Rubin [20], we shall develop a sequence of admissible matrices that have the maximum-likelihood estimate of K subject to this circulant extension as a limit point. The approach parallels that of Miller and Snyder [21] for estimating the power spectrum of a time-series from a single set of data. An important benefit of introducing the periodic extension and using the expectation-maximization algorithm is that estimates of both the scattering function and the reflectance process are obtained simultaneously and can be readily viewed as target images in range and cross-range coordinates; thus, the procedure proposed may be considered to be natural for the imaging problem because both types of images considered separately in the past are obtained directly. As a final comment regarding our use of a circulant extension for K, we note that in estimating a discretized version of the target's scattering function, the class of admissible K is restricted automatically to be those with circulant extensions. For completeness, we also include in the Appendix the equations obtained using the expectation-maximization algorithm for estimating general Toeplitz matrices when the assumption of a circulant extension is not made.

We shall introduce a modification of our notation to indicate that the N samples of the reflectance process are from a stationary periodic-process of period P. Thus, let $b_N(i)$ denote the N-dimensional vector $b(i)$ of (19). We now think of $b_N(i)$ as an N-dimensional subvector of the P-dimensional vector $b_P(i)$ formed from samples of the reflectance process over a full period,

$$b_P(i) = \begin{pmatrix} b(0,m+i) \\ b(1,m+i) \\ \cdot \\ \cdot \\ b(N-1,m+i) \\ \cdot \\ \cdot \\ b(P-1,m+i) \end{pmatrix}. \tag{31}$$

If I_N is the $N\times N$ identity matrix, and if J_R is the $P\times N$ matrix defined by

$$J_R = \begin{pmatrix} I_N \\ 0 \end{pmatrix}, \tag{32}$$

then

$$b_N(i) = J_R^* b_P(i).$$

Also, let b_N denote the NI-dimensional vector b of (18), and let b_P be the PI-dimensional vector with i-*th* block element $b_P(i)$. Then,

$$b_N = M_R^* b_P,$$

where M_R is the $PI \times NI$ block diagonal-matrix

$$M_R = \begin{pmatrix} J_R & 0 & 0 & \cdot & \cdot & 0 \\ 0 & J_R & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & J_R \end{pmatrix}. \tag{33}$$

Furthermore, let $K_N(i)$ denote the $N \times N$ Toeplitz covariance-matrix $K(i)$ of $b_N(i)$ defined in (23), and let $K_P(i)$ denote the $P \times P$ circulant covariance-matrix of $b_P(i)$. Then, the Toeplitz matrix $K_N(i)$ is the upper left block of the circulant matrix $K_P(i)$, as given by

$$K_N(i) = J_R^* K_P(i) J_R.$$

Lastly, let K_P denote the $PI_R \times PI_R$ block-diagonal matrix in the form of (25) with the i-*th* diagonal block being $K_P(i)$. Then, if K_N denotes the $NI_R \times NI_R$ matrix K of (25), there holds

$$K_N = M_R^* K_P M_R.$$

Let W denote the $P \times P$ discrete Fourier-transform matrix scaled so that the columns are orthonormal,

$$W = \frac{1}{\sqrt{P}} \begin{pmatrix} w_P^0 & w_P^0 & \cdot & \cdot & w_P^0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w_P^0 & w_P^k & w_P^{2k} & \cdot & w_P^{(P-1)k} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w_P^0 & w_P^{P-1} & \cdot & \cdot & w_P^{(P-1)(P-1)} \end{pmatrix}, \tag{34}$$

where $w_P = \exp(-j2\pi/P)$. Also, let W_P be the $PI_R \times PI_R$ block-diagonal matrix

$$W_P = \begin{pmatrix} W & 0 & 0 & \cdot & \cdot & 0 \\ 0 & W & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & W \end{pmatrix}. \tag{35}$$

Then, b_P can be represented in rotated coordinates according to

$$a_P = W_P b_P = \begin{pmatrix} a(0) \\ a(1) \\ \cdot \\ \cdot \\ a(I_R - 1) \end{pmatrix}. \tag{36}$$

where $a(i) = Wb_P(i)$. The assumption that $b_P(i)$ originates from a periodic process implies that the PI_R-dimensional vector a_P is normally distributed with zero mean and diagonalized covariance

$$A_P = E\left(a_P a_P^+\right) = W_P K_P W_P^+. \tag{37}$$

We will denote the $(p+iI_R)$-*th* diagonal element of A_P by $\sigma_p^2(i)$; this is the p-*th* diagonal element of the $P \times P$ diagonal matrix $E[a(i)a^+(i)]$.

Let $S(\nu,i)$ be discretized in frequency with P samples taken for $0 \le \nu < 1$. These samples may be obtained from (13) as

$$S\left(\frac{p}{P}, i\right) = \sum_{n=0}^{P-1} K_P(n,i) \exp\left(-j\frac{2\pi n p}{P}\right), \tag{38}$$

for $p = 0, 1, \ldots, P-1$. The p-*th* such sample is just the p-*th* entry in the vector

$$\sqrt{P}\, W K_P(i) e, \tag{39}$$

where e is the P-dimensional unit vector

$$e = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

Substituting (37) into (39), we get

$$\sqrt{P}\,W K_p(i) e = \sqrt{P}\,A(i) W e\,. \tag{40}$$

But $P^{1/2}We$ is a P-dimensional vector of ones and, therefore,

$$S\left(\frac{p}{P}, i\right) = \sigma_p^2(i)\,, \tag{41}$$

which, according to the above definition, is the $(p+iI_R)$-th diagonal element of the diagonal matrix A_P. The entries of the diagonal matrix A_P in (37) are then samples of the scattering function.

The constraint from Section 2 that the scattering function $S(\nu,i)$ be nonzero only for $|\nu| \leq f_{max}\Delta t = \nu_{max}$, for values of ν between $-1/2$ and $+1/2$, may be incorporated at this point in the development. Since A_P is a diagonal matrix of samples of the scattering function, we restrict A_P to have nonzero values only in its top left and bottom right corners. More precisely, let I_{CR} be the smallest odd integer satisfying

$$I_{CR} \geq 2\nu_{max} P\,.$$

I_{CR} is the number of cross-range resolution cells implied by P and ν_{max}. Then, let J_{CR} be the $I_{CR} \times P$ matrix

$$J_{CR} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & 0 & I_2 \end{pmatrix},$$

where I_1 is an $[(I_{CR}+1)/2] \times [(I_{CR}+1)/2]$ identity matrix, and I_2 is an $[(I_{CR}-1)/2] \times [(I_{CR}-1)/2]$ identity matrix. Let M_{CR} be the $I_{CR}I_R \times PI_R$ matrix

$$M_{CR} = \begin{pmatrix} J_{CR} & 0 & 0 & \cdot & \cdot & 0 \\ 0 & J_{CR} & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & J_{CR} \end{pmatrix}.$$

Define Σ_P to be the diagonal matrix

$$\Sigma_P = M_{CR} A_P M_{CR}^*\,.$$

The diagonal elements of Σ_P are the potentially nonzero diagonal elements of A_P. Recognizing that some elements of the diagonal matrix A_P are zero, and using the definition of M_{CR}, we conclude also that

$$A_P = M_{CR}^* \Sigma_P M_{CR}.$$

Then, the set Ω referred to in the definition in Section 2 can now be specified. We restrict consideration to those covariance matrices generated by Σ_P from above, so

$$\Omega = \{ K \in \mathrm{K} : K = M_R^* W_P^* M_{CR}^* \Sigma_P M_{CR} W_P M_R \}.$$

For use with the expectation-maximization algorithm, we identify the *complete data* as (c_P, w), where w is the N-dimensional noise vector defined in (15) and c_P is defined by

$$c_P = M_{CR} a_P.$$

Since elements of a_P corresponding to the zero diagonal elements of A_P are almost surely zero, we see also that

$$a_P = M_{CR}^* c_P.$$

Using this fact, we note from (14) and (20) that the incomplete data r can be obtained from the complete data according to the mapping

$$r = \Gamma^* c_P + w, \tag{42}$$

where we define the following $I_R I_{CR} \times N$ matrix, which will appear throughout the development that follows:

$$\Gamma = \sqrt{2E_T}\, M_{CR} W_P M_R S.$$

The loglikelihood function $L_{cd}(\Sigma_P)$ of the complete data as a function of Σ_P, the diagonal covariance-matrix of c_P, is given by

$$\begin{aligned} L_{cd}(\Sigma_P) &= -\frac{1}{2}\ln(\det(\Sigma_P)) - \frac{1}{2} c_P^* \Sigma_P^{-1} c_P \\ &= -\sum_{i=0}^{I_R-1} \sum_{p=0}^{I_{CR}-1} \ln(\sigma_p(i)) - \frac{1}{2} \sum_{i=0}^{I_R-1} \sum_{p=0}^{I_{CR}-1} |c_p(i)|^2 \sigma_p^{-2}(i), \end{aligned} \tag{43}$$

where all terms that are not a function of Σ_P have been suppressed, and where $c_p(i)$ is the p-*th* element of the I_{CR}-dimensional vector $c_P(i) = J_{CR}^+ a(i)$.

The expectation-maximization algorithm for estimating the covariance of the reflectance process K_P from the incomplete data r is an alternating maximization procedure in which a sequence of estimates of Σ_P having increasing likelihood is obtained first. If $\Sigma_P^{(k)}$ denotes the estimate of Σ_P at stage k, then there is a corresponding element,

$$K_P^{(k)} = W_P^* M_{CR}^* \Sigma_P^{(k)} M_{CR} W_P$$

in a sequence of estimates of K_P having increasing likelihood. Likewise, to the k-*th* element $K_P^{(k)}$ of the sequence of estimates of K_P, there is a corresponding element,

$$K_N^{(k)} = M_R^* K_P^{(k)} M_R,$$

in a sequence of estimates of K_N having increasing likelihood.

Each iteration stage of the expectation-maximization algorithm has an expectation (E) step and a maximization (M) step that must be performed to get to the next step. The E-step requires evaluation of the conditional expectation of the complete-data loglikelihood (43) given the incomplete data r and assuming that the covariance defining the complete data is $\Sigma_P^{(k)}$,

$$Q[\Sigma_P | \Sigma_P^{(k)}] = \mathrm{E}[L_{cd}(\Sigma_P) | r, \Sigma_P^{(k)}]. \tag{44}$$

From (43), we have that

$$Q[\Sigma_P | \Sigma_P^{(k)}] = -\sum_{i=0}^{I_R-1} \sum_{p=0}^{I_{CR}-1} \ln(\sigma_p(i)) - \frac{1}{2} \sum_{i=0}^{I_R-1} \sum_{p=0}^{I_{CR}-1} \mathrm{E}[|c_p(i)|^2 | r, \Sigma_P^{(k)}] \sigma_p^{-2}(i). \tag{45}$$

The M-step yields the estimate $\Sigma_P^{(k+1)}$ at stage $k+1$ as the choice of Σ_P that maximizes this conditional expectation,

$$\Sigma_P^{(k+1)} = \arg\max\, [Q(\Sigma_P | \Sigma_P^{(k)})], \tag{46}$$

subject to the constraint that the maximizer be a diagonal covariance-matrix. From (45), this maximization yields the diagonal matrix $\Sigma_P^{(k+1)}$ with $(p+iI_R)$-*th* diagonal element

$$(\sigma_p^2(i))^{(k+1)} = \mathrm{E}[\,|c_p(i)|^2 | r, \Sigma_P^{(k)}]. \tag{47}$$

Thus, we may write $\Sigma_P^{(k+1)}$ as

$$\Sigma_P^{(k+1)} \overset{d}{=} \mathrm{E}[c_P c_P^* | r, \Sigma_P^{(k)}], \tag{48}$$

where the "d" over the equal sign means that the diagonal terms in the matrix on the left side equal the diagonal terms in the matrix on the right side and that all the off diagonal elements on the left side are zero.

The above expression (48) appears to be complicated because of the several matrices we have defined, but it produces a sequence of covariance estimates having a straightforward interpretation. If we form the matrix $K_P^{(k+1)}$ according to

$$K_P^{(k+1)} = W_P^* M_{CR}^* \Sigma_P^{(k+1)} M_{CR} W_P, \tag{49}$$

we then find that

$$K_P^{(k+1)} = \begin{pmatrix} K_P^{(k+1)}(0) & 0 & 0 & \cdot & \cdot & 0 \\ 0 & K^{(k+1)}(1) & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & K_P^{(k+1)}(I_R-1) \end{pmatrix}, \tag{50}$$

where $K_P^{(k+1)}(i)$ is a $P \times P$ circulant matrix interpreted as the estimate at stage $k+1$ of the covariance $K_P(i)$ of the P-periodic reflectance process at delay $m+i$. Miller and Snyder [21] show that the (n,m)-element of this circulant matrix is given by

$$\frac{1}{P} \sum_{p=0}^{P-1} \mathrm{E}[\, b(p,i) b^*\left(<p+m-n>_P, i\right) \mid r, K_P^{(k)}], \tag{51}$$

where $<a>_P = a \bmod P$. Equation (51) has an intuitively appealing form. If the reflectivity process $b(n,i)$ could be observed for all instants $n = 0, 1, \ldots, P-1$ in a period and for each i independently, then the maximum-likelihood estimate of the covariance $K_P(i)$ would be the arithmetic average of the lagged products

$$\frac{1}{P} \sum_{p=0}^{P-1} b(p,i) b^*\left(<p+m-n>_P, i\right). \tag{52}$$

Equations (50) and (51) indicate that one should simply substitute the conditional mean estimate of an unknown lagged product into this expression to form the maximum-likelihood estimate of the covariance when only the incomplete data are known.

estimating Σ_P and K_P

The maximum-likelihood estimate of Σ_P is a limit point of the sequence defined in (48). The terms on the right side of this equation can be evaluated as follows. Let the conditional-mean estimate of c_P in terms of the incomplete data r be defined at stage k by

$$\hat{c}_P^{(k)} = E[\, c_P \mid r, \Sigma_P^{(k)}]. \tag{53}$$

Then, (48) can be rewritten in the form

$$\Sigma_P^{(k+1)} \stackrel{d}{=} E[\left(c_P - \hat{c}_P^{(k)}\right)\left(c_P - \hat{c}_P^{(k)}\right)^* \mid r, \Sigma_P^{(k)}] + \hat{c}_P^{(k)} \hat{c}_P^{(k)*} . \tag{54}$$

Now examination of (42) shows that forming the conditional-mean estimate (53) of c_P from r is a standard problem in linear estimation-theory. From Tretter [25, Ch. 14], for example, we find that

$$\hat{c}_P^{(k)} = \Sigma_P^{(k)} \Gamma [\Gamma^* \Sigma_P^{(k)} \Gamma + N_0 I]^{-1} r . \tag{55}$$

Furthermore, the first term on the right side in (54) is the covariance of the estimation error when c_P is estimated from r. Also from Tretter [25, Ch. 14], we have

$$\begin{aligned} E[& \left(c_P - \hat{c}_P^{(k)}\right)\left(c_P - \hat{c}_P^{(k)}\right)^* \mid r, \Sigma_P^{(k)}] \\ & = \Sigma_P^{(k)} - \Sigma_P^{(k)} \Gamma [\Gamma^* \Sigma_P^{(k)} \Gamma + N_0 I]^{-1} \Gamma^* \Sigma_P^{(k)} . \end{aligned} \tag{56}$$

In summary, the following steps are performed to produce a sequence $\Sigma_P^{(0)}$, $\Sigma_P^{(1)}$, ..., $\Sigma_P^{(k)}$, ... of estimates of Σ_P for which the corresponding sequence of likelihoods is nondecreasing:

1. set $k = 0$, select a starting estimate $\Sigma_P^{(0)}$;
2. calculate the estimate of c_P according to (55);
3. calculate the error covariance according to (56);
4. update the estimate of Σ_P according to (54);
5. if "last iteration" then stop, else replace k by $k+1$ and go to 2.

The starting value in step 1 can be any positive-definite, diagonal covariance-matrix of dimension $I_R I_{CR} \times I_R I_{CR}$.

A sequence of estimates of K_P having increasing likelihood is obtained from the sequence of estimates of Σ_P according to (49).

forming the scattering-function image

From (41), the diagonal elements of the $I_R I_{CR} \times I_R I_{CR}$-dimensional matrix Σ_P are sample values of the scattering function, with the scattering function samples at delay $m+i$ given by the I_{CR} diagonal elements of the i-*th* $I_{CR} \times I_{CR}$-dimensional diagonal block of Σ_P. We may, therefore, simply regard $\Sigma_P^{(k)}$ as the stage k estimate of the scattering function. The stage-k scattering-function image of the target in range (i coordinate) and cross range (p coordinate) can be displayed as follows. Let $\Sigma_P^{(k)}(i)$ denote the i-th $I_{CR} \times I_{CR}$-dimensional diagonal block of $\Sigma_P^{(k)}$, and denote the p-th diagonal element of $\Sigma_P^{(k)}(i)$ by $s(p,i) = [\sigma_p^2(i)]^{(k)}$ for $p = 0, 1, \ldots I_{CR}-1$. Then, $s(0,i)$ is displayed at range $m+i$ and cross range corresponding to a doppler shift of zero; $s(1,i)$ and $s(I_{CR}-1,i)$ are displayed at range $m+i$ and cross range corresponding to a doppler shift of $\nu = 1/P$ and $\nu = -1/P$, respectively; $s(2,i)$ and $s(I_{CR}-2,i)$ are displayed at range $m+i$ and cross range corresponding to a doppler shift of $\nu = 2/P$ and $\nu = -2/P$, respectively; and so forth, with $s(p,i)$ and $s(I_{CR}-p,i)$ displayed at range $m+i$ and cross range corresponding to a doppler shift of $\pm p/P$ for $p = 1,2,\ldots, (I_{CR}-1)/2$ when I_{CR} is odd.

forming the reflectance-process image

It is interesting to note that the k-*th* stage conditional-mean estimate of c_P, given the measurements r and assuming that the second-order statistics of reflectance are given by the k-*th* stage estimate of the scattering function, is used to form the estimate of Σ_P at stage $k+1$ when the expectation-maximization algorithm is used. This estimate is of very much interest in its own right because, from (36) and its definition, the I_{CR} elements of $c_P(i)$ are the potentially nonzero Fourier-transform coefficients of the reflectance process $h_P(i)$. The target's reflectance image at stage k is formed by placing these elements at range $m+i$ and cross range in the same manner as described above for the scattering-function image.

convergence issues

There are some important properties of the iteration sequence (48) which are worth mentioning. First, each step is in an improving direction. This is shown by writing (54) out as

$$\Sigma_P^{(k+1)} = \Sigma_P^{(k)} + \Sigma_P^{(k)} \Theta^{(k)} \Sigma_P^{(k)} . \tag{57}$$

where

$$\Theta^{(k)} = \Gamma K_r^{(k)-1}\left(rr^* - K_r^{(k)}\right)K_r^{(k)-1}\Gamma^*, \tag{58}$$

and where

$$K_r^{(k)} = \Gamma^* \Sigma_P^{(k)}\Gamma + N_0 I \tag{59}$$

is the k-*th* estimate of the covariance K_r of r. Next, the trace condition (27) which the maximum-likelihood estimate must satisfy is reexamined. From the assumption of the P-periodicity of the reflectance process and the matrix definitions given, the admissible variations δK must be of the form

$$\delta K = M_R^* W_P^* M_{CR}^* \delta\Sigma M_{CR} W_P M_R . \tag{60}$$

Here, $\delta\Sigma$ is a diagonal matrix of the same dimensions as Σ. The trace condition (27) then becomes

$$2E_T^{-1}\operatorname{Tr}\left(K_r^{-1}\left(\Gamma^*\Sigma_P\Gamma + N_0 I - rr^*\right)K_r^{-1}\Gamma^*\delta\Sigma\Gamma\right) = 0 . \tag{61}$$

Using the fact that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ and evaluating this trace at the k-*th* iterate, we see that the trace on the left side of (61) is equal to

$$2E_T^{-1}\operatorname{Tr}\left(\Theta^{(k)}\delta\Sigma\right). \tag{62}$$

According to (57), $\Sigma_P^{(k)}$ is changed at each stage by adding the diagonal elements of

$$\Sigma_P^{(k)}\Theta^{(k)}\Sigma_P^{(k)} \tag{63}$$

to $\Sigma_P^{(k)}$. Define

$$\delta\Sigma^{(k)} \stackrel{d}{=} \Sigma_P^{(k)}\Theta^{(k)}\Sigma_P^{(k)}. \tag{64}$$

as these diagonal elements. Then, evaluating the trace at this variation gives

$$\operatorname{Tr}\left(\Theta^{(k)}\delta\Sigma^{(k)}\right) \geq 0 . \tag{65}$$

This shows that the variation

$$\delta\Sigma^{(k)}$$

is in an improving direction. Furthermore, we are guaranteed that the incomplete-data loglikelihood is nondecreasing as a result of the M-step of the expectation-maximization

algorithm because at this step, the conditional expectation of the complete-data loglikelihood given the incomplete data and the last iterate for Σ_P is maximized over Σ_P. As shown in [20] and [21], this implies that the incomplete-data loglikelihood is nondecreasing.

Lemma 3. Assume that $N_0 > 0$ and $\Sigma_P^{(0)}$ is positive definite. Then:

(i), $\Sigma_P^{(k)}$ is positive definite for all k; and (ii), all stable points satisfy the trace condition (27) for all admissible variations (60).

The proof of the first part of Lemma 3 is in the Appendix. For the second part, since the diagonal elements of $\Sigma_P^{(k)}$ are positive, (65) holds with equality if and only if the diagonal elements of $\Theta^{(k)}$ are zero. Notice that if $\Sigma_P^{(k+1)} = \Sigma_P^{(k)}$, then the diagonal elements (64) are zero. This implies that the diagonal elements of $\Theta^{(k)}$ are zero and, hence, that

$$\mathrm{Tr}\left(\Theta^{(k)}\delta\Sigma\right)=0 \tag{66}$$

for all diagonal $\delta\Sigma$. Thus, all stable points satisfy the trace condition (27) for admissible variations.

computational considerations

The computations required to produce radar images with our method are specified by (54), (55), and (56). The number of iterations of these equations that are required to produce an image near the convergence point is presently unknown. Our experience in using an iterative algorithm to produce maximum-likelihood images for emission tomography suggests that 50 to 100 iterations may be necessary, but this is only a guess that will not be verified until some experiments are completed. Some form of specialized processor to accomplish each iteration stage efficiently will probably be needed to produce images in practically useful times. One possible approach is the following. The matrix product

$$\Gamma = \sqrt{2E_T}\,M_{CR}W_P M_R S$$

is required at each iteration stage and does not change. This $I_R I_{CR}\times N$-dimensional matrix can, therefore, be computed once off line, stored, and then used as needed during on-line computations. Then, at iteration stage k, the following on-line computations can be performed:

1. compute the $N\times N$-dimensional matrix A defined by $A = \Gamma^{+}\Sigma_P^{(k)}\Gamma + N_0 I$;
2. compute the $I_R I_{CR}\times N$-dimensional matrix B defined by $B = \Sigma_P^{(k)}\Gamma$;

3. compute $BA^{-1}r$ and the diagonal elements of $\Sigma_P^{(k)} - BA^{-1}B^{+}$.

The computations in 3 can be accomplished in about $4N+I_R I_{CR}-2$ time steps using the systolic array described by Comon and Robert [26] augmented, as they suggest, by one row to accomplish the postmultiplication of BA^{-1} by r and by $I_R I_{CR}$ rows to accomplish the postmultiplication by B^{+}. The matrix multiplications in 1 and 2 for determining A and B can also be performed rapidly on a systolic array. More study of implementation approaches is needed, but it does not appear that the computational complexity of our new imaging algorithm needs to be a limitation to its practical use.

The choice of N, I_R and I_{CR} is important for the computations. These parameters are selected to achieve a desired range and cross-range resolution and are, therefore, problem dependent, but the same considerations used with other approaches to radar imaging can be used in selecting them. Choosing the product $I_C I_{CR}$ to be on the order of N will, in some sense, make the imaging problem well defined because the number of unknown parameters $I_C I_{CR}$ that need to be estimated to form the image is then comparable to the number of measurements N. On the other hand, the choice of P is unique to our approach. As stated, we need $P \geq N$, but no upper limit is given. In [24], it is shown that as P increases towards infinity so does the maximum value of the incomplete-data loglikelihood function, with probability one. Thus, P cannot be made arbitrarily large from a theoretical standpoint. Any computation involving a matrix with one dimension equal to P can be performed off line.

4. Conclusions

The expressions we have obtained for forming images of diffuse, fluctuating radar targets are based on the model stated in Section 2. The target reflectance is assumed to introduce wide-sense-stationary uncorrelated-scattering (WSSUS) of the transmitted signal with no glint or specular components being present. The reflectance process is assumed to be a WSSUS Gaussian process with unknown second-order statistics given by a delay-dependent covariance or scattering function. Echos of the transmitted signal are received from all the reflecting patches that make up the target, with each patch introducing some propagation delay, doppler shift, and random amplitude-scaling into the signal it reflects. The superposition of the echos from all the patches is received in additive noise. Thus, the reflectance process is only observed indirectly following a linear superposition and in additive noise, so neither the reflectance process nor its second-order statistics are known. Target images are made by displaying estimates of either the reflectance process or its second-order statistics (scattering function) based on processing the received signal. In Section 2, we derived the trace condition which the maximum-likelihood estimate of the covariance of the reflectance must satisfy, and we concluded that this condition is too complicated to solve explicitly for the estimate. This motivated the introduction in Section 3 of the incomplete-complete data model and the use of the expectation-maximization algorithm, which results in a sequence of estimates of the scattering function having increasing likelihood. A corresponding sequence of estimates of the reflectance process is also obtained.

There are a number of issues yet to be resolved for the approach to radar imaging which we have presented. One of the most important is resolving how glint and specular components in the return echos should be modeled and accommodated in the formation of the images. The selection of transmitted signals to produce good images is an important subject about which little study has been made. The quality of target images obtained with our new approach is not known at present; to study this issue, we are presently implementing a computer simulation so that comparisons to alternative processing strategies can be made. The equations we have developed are computationally demanding, so special processing architectures will be important to make their use practical.

5. Acknowledgment

We are grateful to Dr. Richard E. Blahut of I.B.M., Owego NY for reading a draft of this paper and making several helpful suggestions for improving it and to Dr. Jeffrey Shapiro of M.I.T., Cambridge, MA for bringing Reference [9] to our attention.

6. References

1. D. R. Wehner, *High Resolution Radar*, Artech House, Dedham, MA, 1987.
2. D. L. Mensa, *High-Resolution Radar Imaging*, Artech House, Dedham, MA, 1984.
3. M. Bernfeld, "CHIRP Doppler Radar," Proceedings of the IEEE, Vol. 72, pp. 540-541, April 1984.
4. D. L. Snyder, H. J. Whitehouse, J. T. Wohlschlaeger, and R. Lewis, "A New Approach to Radar/Sonar Imaging," Proc. SPIE Conference on Advanced Algorithms and Architectures, Vol. 696, San Diego, CA, August 1986.
5. J. L. Walker, "Range-Doppler Imaging of Rotating Objects," IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-16, No. 1, pp. 23-52, January 1980.
6. J. Shapiro, B. A. Capron, and R. C. Harney, "Imaging and Target Detection with a Hetrodyne-Reception Optical Radar," Applied Optics, Vol. 20, No. 19, pp. 3292-3313, October 1981.
7. H. L. Van Trees, *Estimation, Detection, and Modulation Theory, Vol. 3*, John Wiley and Sons, NY, 1971.
8. V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, "A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-4, No. 2, pp. 643-652, March 1982.
9. N. T. Gaarder, "Scattering Function Estimation," IEEE Trans. on Information Theory, Vol. IT-14, No. 5, pp. 684-693, September 1968.
10. P. E. Green, "Radar Measurements of Target Scattering Properties," Radar Astronomy, V. V. Evans and T. Hagfors, Eds., McGraw-Hill, New York, Ch. 1, pp. 1-78, 1968.
11. T. Kailath, "Measurements in Time-Variant Communication Channels," IRE Trans. on Information Theory, Vol. IT-8, pp. 829-8236, September 1962.
12. R. G. Gallager, "Characterization and Measurement of Time-and-Frequency Spread Channels," M.I.T. Lincoln Laboratory Tech. Rpt. 352, Lexington, MA, April 1964.
13. T. Hagfors, "Some Properties of Radio Waves Reflected From the Moon and Their Relation to the Lunar Surface," J. Geophys. Res., Vol. 66, p. 777, 1961.
14. T. Hagfors, "Measurement of Properties of Spread Channels by the Two-Frequency Method with Application to Radar Astronomy," M.I.T. Lincoln Laboratory Tech. Rpt. 372, Lexington, MA, January 1965.
15. R. Price, "Maximum-Likelihood Estimation of the Correlation Function of a Threshold Signal and Its Application to the Measurement of the Target Scattering Function in Radar Astronomy," M.I.T. Lincoln Laboratory Group Rpt. 34-G-J, Lexington, MA, May 1962.
16. M. V. Levin, "Estimation of the Second-Order Statistics of Randomly Time-Varying Linear Systems," M.I.T. Lincoln Laboratory Group Rpt. 34-G-7, Lexington, MA, November 1962.
17. L. G. Abraham, *et al.*, "Tropospheric-Scatter Propagation Tests Using a RAKE Receiver," 1965 IEEE Communications Convention Record, pp. 685-690, 1965.
18. B. M. Sifford, *et al.*, "HF Time- and Frequency-Dispersion Effects: Experimental Validation of an FSK Error Rate Model," Stanford Research Institute Tech. Rpt. 4, Menlo Park, CA, March 1965.
19. B. Reiffen, "On the Measurement of Atmospheric Multipath and Doppler Spread by Passive Means," M.I.T. Lincoln Laboratory Tech. Note 1965-6 G-66, March 1965.
20. A. D. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data Via the EM Algorithm," J. of the Royal Statistical Society, Vol. B, 39, pp. 1-37, 1977.

21. M. I. Miller and D. L. Snyder, "The Role of Likelihood and Entropy in Incomplete-Data Problems: Applications to Estimating Point-Process Intensities and Toeplitz-Constrained Covariances," Proceedings of the IEEE, Vol. 75, No. 7, July 1987.

22. M. Turmon and M. I. Miller, "Performance Evaluation of Maximum-Likelihood Estimates of Toeplitz Convariances Generated with the Expectation-Maximization Algorithm," 1987 Conference on Information Sciences and Systems, Johns Hopkins University, Baltimore, MD, pp. 25-27, 1987.

23. J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of Structured Covariance Matrices," Proc. IEEE, Vol. 70, No. 9, pp. 963-974, September 1982.

24. D. R. Fuhrmann and M. I. Miller, "On the Existence of Positive Definite Maximum-Likelihood Estimates of Structured Covariance Matrices," IEEE Transactions on Information Theory, in review.

25. S. A. Tretter, *Introduction to Discrete-Time Signal Processing*, John Wiley and Sons, New York, 1976.

26. P. Comon and Y. Robert, "A Systolic Array for Computing BA^{-1}," IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-35, No. 6, pp. 717-723, June 1987.

7. Appendix

proof of Lemma 1.

From the definition of the loglikelihood function in (26), we have

$$
\begin{aligned}
&\frac{1}{\alpha}\left(L_{id}(K+\alpha\delta K)-L_{id}(K)\right) \\
&\quad = -\frac{1}{2\alpha}r^*\left(\left(K_r+\alpha 2E_T S^*\delta K S\right)^{-1}-K_r^{-1}\right)r-\frac{1}{2\alpha}\ln\left(\det\left(K_r+\alpha 2E_T S^*\delta K S\right)\det K_r^{-1}\right).
\end{aligned}
\tag{A1}
$$

where K_r is the covariance of the incomplete data r as given in (26). Examining the first term on the right, we have that it equals

$$
\begin{aligned}
&-\frac{1}{2\alpha}r^*K_r^{-1}\left(\left(I+\alpha 2E_T S^*\delta K S K_r^{-1}\right)^{-1}-I\right)r \\
&\qquad = \frac{1}{2\alpha}r^*K_r^{-1}\left(\alpha 2E_T S^*\delta K S K_r^{-1}\right)\left(I+\alpha 2E_T S^*\delta K S K_r^{-1}\right)^{-1}r \\
&\qquad = \frac{1}{2}r^*K_r^{-1}2E_T S^*\delta K S K_r^{-1}r+O(\alpha).
\end{aligned}
\tag{A2}
$$

Examining the second term on the right in (A1), we have

$$
\begin{aligned}
-\frac{1}{2\alpha}\ln\left(\det\left(I+\alpha 2E_T S^*\delta K S K_r^{-1}\right)\right) &= -\frac{1}{2\alpha}\ln(\det(I+\alpha B)) \\
&= -\frac{1}{2\alpha}\ln\left(1+\alpha\operatorname{Tr}(B)+\ldots+\alpha^n\det(B)\right) \\
&= -\frac{1}{2}\operatorname{Tr}(B)+O(\alpha),
\end{aligned}
\tag{A3}
$$

where B is defined in the first equality. For any $K \in \Omega$ to be a local maximum, a small variation in K in an admissible direction cannot increase $L_{id}(K)$, or

$$
\lim_{\alpha\to 0^+}\frac{1}{\alpha}\left(L_{id}(K+\alpha\delta K)-L_{id}(K)\right)\le 0
\tag{A4}
$$

for all admissible variations δK. If K is an interior point, then $-\delta K$ is also an admissible variation and (A4) becomes an equality. Substituting (A2) and (A3), we get

$$r^* K_r^{-1} 2E_T S^* \delta K S K_r^{-1} r - \mathrm{Tr}\left(2E_T S^* \delta K S K_r^{-1}\right) = 0, \tag{A5}$$

which is the trace condition (27).

proof of Lemma 2

Suppose that K satisfies (27) and (29) for all admissible variations δK. We now show that (29) is simply the second derivative of $L_{id}(K)$ in the direction δK by taking the limit

$$\begin{aligned}
&\lim_{\alpha \to 0} \frac{1}{|\alpha|} \mathrm{Tr}\Big(\big(2E_T S^*(K + \alpha\delta K)S + N_0 I\big)^{-1}\big(rr^* - N_0 I - 2E_T S^*(K + \alpha\delta K)S\big) \\
&\qquad \times \big(2E_T S^* K S + \alpha 2E_T S^* \delta K S + N_0 I\big)^{-1} S^* \delta K S \\
&\qquad - K_r^{-1}\big(rr^* - N_0 I - 2E_T S^* K S\big) K_r^{-1} S^* \delta K S\Big) \\
&= \mathrm{Tr}\left(K_r^{-1} 2E_T S^* \delta K S K_r^{-1}\big(2E_T S^* K S + N_0 I - 2rr^*\big) K_r^{-1} 2E_T S^* \delta K S\right).
\end{aligned} \tag{A6}$$

Thus, the conditions of Lemma 2 are the standard sufficient conditions for a point K to be the local maximum of $L_{id}(K)$. Equation (27) says that the first derivative of $L_{id}(K)$ is zero at K. Equation (29) says that the second derivative is negative definite along admissible variations from K. A necessary condition for K to be a relative maximum is that this last expression evaluated at K be equal to or less than zero for all admissible variations δK. Under the assumptions in Section 4, admissible variations are given by (60). Substituting (60) into (A5) and evaluating for all diagonal matrices $\delta\Sigma$ gives the second-order necessary condition.

estimating a general Toeplitz matrix

In Section 4, we derived a sequence of estimates for a covariance matrix subject to the constraint that the estimates must be circulant Toeplitz matrices. For completeness, we develop and discuss in this section the equations for estimating a covariance matrix subject to the weaker constraint that the estimates be general Toeplitz matrices. Similar equations for other constraints on the Toeplitz matrices are easily obtained by mimicking the steps in the main body of this paper.

Let the complete data be (b, w), and let b be normally distributed with zero mean and covariance K, as given in (27). The complete-data loglikelihood is

$$L_{cd}(K_P) = -\frac{1}{2}\ln \, det(K_P) - \frac{1}{2}b_P^* K_P^{-1} b_P, \tag{A7}$$

where all terms that are not a function of K_P have been suppressed. Maximizing this function gives the trace condition

$$\mathrm{Tr}\left(K^{-1}\left(bb^* - K\right)K^{-1}\delta K\right) = 0, \tag{A8}$$

which the maximum-likelihood estimate K must satisfy. Performing the E and M steps of the expectation-maximization algorithm yields the following iteration sequence for the elements $K(n,i)$, $n = 0, 1, \ldots, N-1$, of the covariance matrix $K(i)$ defined in (23):

$$K^{(k+1)}(n,i) = \frac{1}{N-n}\mathrm{E}\left[\sum_{j=0}^{N-n-1} b(j,m+i)b^*(j+n,m+i) \,|\, r, K^{(k)}\right]. \tag{A9}$$

In matrix form,

$$K^{(k+1)} = K^{(k)} + 2E_T K^{(k)} S \, K_r^{(k)-1}\left(rr^* - 2E_T S^* K^{(k)} S - N_0 I\right)K_r^{(k)-1} S^* K^{(k)}, \tag{A10}$$

where

$$K_r^{(k)} = 2E_T S^* K^{(k)} S + N_0 I. \tag{A11}$$

If this iteration converges to a stable point, then the trace condition is satisfied at that point, as may be shown by using the same arguments as in Section 4. It is worth restating that the reason this iteration is not recommended here is that the probability that the iteration sequence generates a singular estimate for K approaches one as N gets large. By restricting consideration to Toeplitz matrices with circulant extensions, the loglikelihood function is bounded with probability one for finite extensions and a positive definite K is generated with probability one, as proven by Fuhrmann and Miller [24].

proof of Lemma 3, part (i)

Assume that the initial guess $\Sigma_P^{(0)}$ for Σ_P is positive definite and that $N_0 > 0$. We will now show that if $\Sigma_P^{(k)}$ is positive definite, then so is $\Sigma_P^{(k+1)}$, and thus, by induction, $\Sigma_P^{(k)}$ is positive definite for all k. One key to following this derivation is the matrix identity

$$B(I + AB)^{-1} = (I + BA)^{-1}B. \tag{A12}$$

This identity is used to rewrite (57) as

$$\begin{aligned}\Sigma_P^{(k+1)} &\overset{d}{=} H(N_0\Sigma_P^{(k)}\Gamma\Gamma^*\Sigma_P^{(k)} + N_0^2\Sigma_P^{(k)} \\ &\quad + \Sigma_P^{(k)}\Gamma r r^*\Gamma^*\Sigma_P^{(k)})H^* \\ &= N_0\left(H\Sigma_P^{(k)}\Gamma\right)\left(\Gamma^*\Sigma_P^{(k)}H^*\right) \\ &\quad + \left(H\Sigma_P^{(k)}\Gamma r\right)\left(r^*\Gamma^*\Sigma_P^{(k)}H^*\right) \\ &\quad + N_0^2 H\Sigma_P^{(k)}H^*,\end{aligned} \tag{A13}$$

where we have defined H according to

$$H = \left(\Sigma_P^{(k)}\Gamma\Gamma^* + N_0 I\right)^{-1}. \tag{A14}$$

Clearly, all the diagonal elements of (A13) are greater than or equal to zero. To show that they are strictly positive, we look at the last term and get that the i-*th* diagonal element is

$$\begin{aligned}\left(N_0^2 H\Sigma_P^{(k)}H^*\right)_{ii} &= N_0^2\sum_{j=0}^{I_R I_{CR}-1}(H)_{ij}\left(\Sigma_P^{(k)}\right)_{jj}\left(H^*\right)_{ji} \\ &= N_0^2\sum_{j=0}^{I_R I_{CR}-1}\left|(H)_{ij}\right|^2\left(\Sigma_P^{(k)}\right)_{jj},\end{aligned} \tag{A15}$$

which is clearly positive when $N_0 > 0$ since H is invertible and all diagonal elements of $\Sigma_P^{(k)}$ are positive.

4.2 Appendix 2. Parameter Values Used in Confidence-Weighted Simulation Experiments

This appendix contains a table updating parameter values given in the first six-month progress report of this contract. These are the values being used in computer-simulation studies of the confidence-weighted algorithm.

We assume the following for the gaussian envelope waveforms:

Quantity	Symbol	Value
Distance to target:	R	12.5 nautical miles = 23150 m
Number of Bursts:	N	128
Number of Pulses/Burst:	n	1
Center Frequency:	fc	3 GHz
Wavelength:	lambda	0.1 m
Speed of Light:	C	3x10^8 m/sec

Given

Target Angular Rotation Rate:	w	
Downrange Resolution:	rd	
Crossrange Resolution:	rc	
Ratio of major to minor axis lengths for desired ambiguity function:	e	

The following are uniquely determined for the burst sequence:

Target Diameter:	D	64*rd
Round Trip Time:	Tr	2*R/C
Delay Resolution:	rt	2*rd/C
Frequency Resolution:	rf	2*w*rc/lambda
Constant 1:	a1	2*pi*e
Constant 2:	a2	2*pi/e

The following are uniquely determined for a given burst, burst (i):

angle major axis makes with the positive delay axis:	angle	(i-1)/n*pi
Parameter 1:	p1	a1*sin(angle)^2 + a2*cos(angle)^2
Parameter 2:	p2	(a2 - a1)*sin(2*angle)
FWHM Burst Duration:	T	sqrt (p1*rt/rf)/(2*pi)
FM Sweep Rate:	b	p2/(8*pi*T^2)
Burst Bandwidth:	BW	b*T
Total Illumination Time:	Tt	sum 3*T(i) + (N-1)*Tr
Total Aspect Change:	W	w*Tt

If we take rc = rd = 0.6 m, e = 1.3e8, w = 8.88x10^-3 radians/sec, we get the following representative values:

Target Diameter:	D	38.4 m
Round Trip Time:	Tr	154 usec
Delay Resolution:	rt	4 nsec
Frequency Resolution:	rf	0.1066 Hz
Constant 1:	a1	2.6e8*pi
Constant 2:	a2	1.5e-8*pi
Minimum FWHM Burst Duration:	Tmin	16.0 nsec
Maximum FWHM Burst Duration:	Tmax	2.07 sec
Minimum FM Sweep Rate:	bmin	15.6 Hz/sec
Maximum FM Sweep Rate:	bmax	+/- 3.41 GHz/sec
Maximum Burst Bandwidth:	BWmax	346. MHz
Total Illumination Time:	Tt	506. sec
Total Aspect Change:	W	257. degrees

END
DATE
FILMED
5-88
DTIC